REPRINT FROM THE

PROCEEDINGS

OF THE

ROYAL SOCIETY OF EDINBURGH.

SESSION 1905–1906.

VOL. XXVI.—PART VI.

Statistical Studies in Immunity: The Theory of an Epidemic.

By John Brownlee, M.D. (Glas.).

EDINBURGH:

Published by ROBERT GRANT & SON, 107 Princes Street, and WILLIAMS & NORGATE, 14 Henrietta Street, Covent Garden, London.

MDCCCCVII.

Price One Shilling and Sixpence.

Statistical Studies in Immunity: The Theory of an Epidemic.

By John Brownlee, M.D. (Glas.). *Communicated by* Dr R. M. Buchanan.

(Read June 18, 1906.)

The rise and decline of epidemics of infectious diseases have been subjects of interest since the earliest times, but the scientific determination of the laws which govern their course offers even yet a wide and almost unworked field. Not but what a large amount of observation has been made on many of the conditions under which epidemics appear and pass away. Many epidemics are seasonal, and these have been studied; but the lack of any means of determining the course which a given epidemic might have taken in the presence of somewhat different conditions has made the deduction of certain conclusions impossible. Even the laws which regulate solitary outbursts of disease, the special subject of this paper, have been little studied. Explanations offered have varied with the period in history. We find that the pestilence which afflicted Israel for David's sin stayed at the threshing-floor of Araunah the Jebusite; we find in the *Iliad* that with the return of Chryseis "the dread clang of Apollo's silver bow" ceased; later, the great fire of London is commonly believed to have exterminated the plague, while at this moment the hasty cleansing of a town by a terrified sanitary authority is by many thought to be the direct cause of the disappearance of an epidemic; but in all these cases there is a misinterpretation of facts, which is due largely to the absence of any real knowledge of the underlying laws.

The most important contribution to the subject from the point of view of this paper is one by the late Dr Farr, who had that genius which permitted him to perceive a large part of the laws which govern progressions of figures. In 1866, when the cattle plague was making most extensive ravages in this country, and when, from the rate of its progress, there seemed no end to the damage it might do, he wrote a letter in which he showed that, as the rate at which the disease was extending was already lessening, the acme and the decline of the epidemic might soon be expected.

Later, he applied the same method to the case of smallpox in 1871–2, and fitted a curve to the latter portion of the epidemic. His description of his method is not clear, but in a paper by Dr G. H. Evans (*Trans. Epid. Soc.*, 1874–5) it is given in detail. The method practically amounts to assuming that the second difference of the logarithms of successive ordinates of an epidemic curve is a constant, and using a value of this constant, deduced from an early portion of the epidemic, to predict the succeeding portion.

The method in the terminology of finite differences is as follows:—

If $u = \log y$ where y is the ordinate of the epidemic curve, then

$$\Delta^2 u = -c \text{ (a constant by Dr Farr's supposition)}$$

of which the integral is

$$u = -\frac{cx^2}{2} + Ax + B$$

or as $\log y = u$

$$y = e^{-\frac{cx^2}{2} + Ax + B}$$

which is the equation to the normal curve of probability.

Dr Farr does not seem to have noticed that the application of his arithmetical law leads to this curve. As a matter of fact, it is a very good approximation to the middle parts of some epidemics, though it does not provide a specially good fit for the whole course of those to which he applied it. In the examples of this method given by Dr Evans nothing more is attempted. The real difficulty in the application consists in finding a good value of the constant from the early portion of the epidemic.

This is all the literature of this subject. My attention was specially drawn to this matter when considering recently some questions of immunity. The interpretation of some of the facts required accurate knowledge of the epidemic processes. I was struck, when I began to examine the course of epidemics, by the close resemblance which many bore to the probability distributions developed by Professor Pearson; and, without any knowledge that Dr Farr had already come indirectly to fit the epidemic curve to that of the normal frequency of error, I applied the methods now in use of fitting probability distributions to statistics.

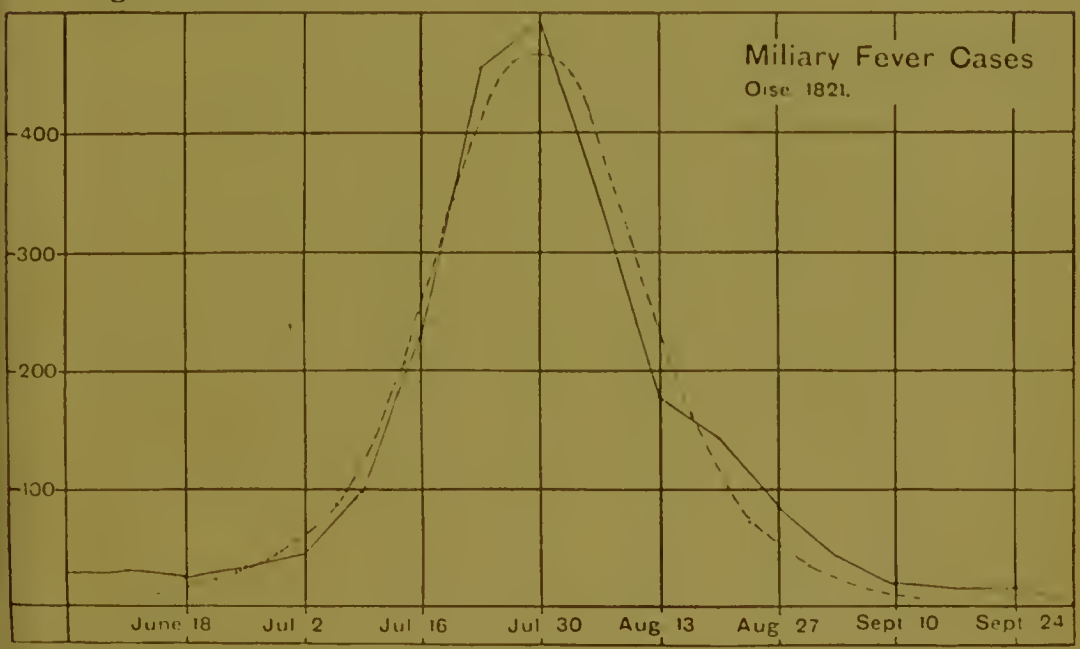
The result of these calculations has been to show that the curve of the solitary epidemic is singularly constant. It almost uniformly corresponds to that probability distribution which Professor Pearson has termed type IV., * so that this curve may be chosen for interpolation, as one which gives a very clear representation of the facts. Even when the epidemic becomes symmetrical, it is to the form of type IV., where ν is equal to zero, that the course approximates, and not to the normal probability curve. The precise limitation of an epidemic in point of time does not alter the form referred to, for in this case also the constants generally indicate a curve of type IV., though the latter is unlimited on either side. As the first example of curve fitting to the course of an epidemic wave, miliary fever (Table A, No. 1) † has been chosen. This instance does not afford an example of a curve which gives a good fit, but it is a disease of which the cause and the means of propagation are absolutely unknown, and therefore one which spreads as near as possible in natural conditions. It affects country districts, so that as new townships are invaded at all stages of the epidemic, and as these tend to become more numerous as the disease extends, a supply of susceptible persons is constantly furnished and a criterion given for the approximate estimation of the infectivity of the organism. It is, in addition, an explosive disease, and thus no difficulty arises about the start and close of the outburst. All the factors which make for the investigation of an epidemic type are thus present in this case. The course of an epidemic of this disease is illustrated in diagram I., along with the theoretical interpolation curve as well. As the number of persons exposed increased from the beginning to the end of the epidemic, it is seen that the decline in the number of cases must be due to the loss of infectivity in the germ itself, and not to the lack of individuals who may be supposed open to the contagion. Equally characteristic examples are afforded by the great plagues of London (Table A, Nos. 2 and 3). That of 1665 is specially interesting, as we have many contemporary accounts of the conditions which obtained. Great migrations from the city began as soon as the plague established itself, but the disease had barely begun to abate when the return

* See Note at end of paper.

† Rayer, *Histoire de l'épidémie de Suette Miliare* en 1821, Paris, 1822.

of emigrants began in great numbers. On this Pepys makes anxious remarks in his Diary, and speculates on the occurrence of a recrudescence of the malady, but the infecting power of the organism was exhausted; and though great numbers of susceptible persons came from the country into the zone of infection, even, it is said, occupying the beds of those who had been afflicted, no further extension of the disease ensued. The curve of this epidemic is given in diagram II. The figures on which it is based are taken partly from the London Bills of Mortality * and partly

Diagram I.



from Pepys' Diary. For comparison, the constants of the curve representing the course of the great outbreak of plague in London in 1563 † are given.

As further examples of the epidemic course, influenza ‡ and cholera (Table A, Nos. 4, 5, 6, and 7) are illustrated (diagrams III., IV., and V.), and the constants given in the table. It is to be noted that the curves are again of the type IV. The epidemics of cholera include that in Exeter in the summer of 1832,§ and

* Quoted also in Creighton's *History of Epidemics in Britain*, vol. ii.

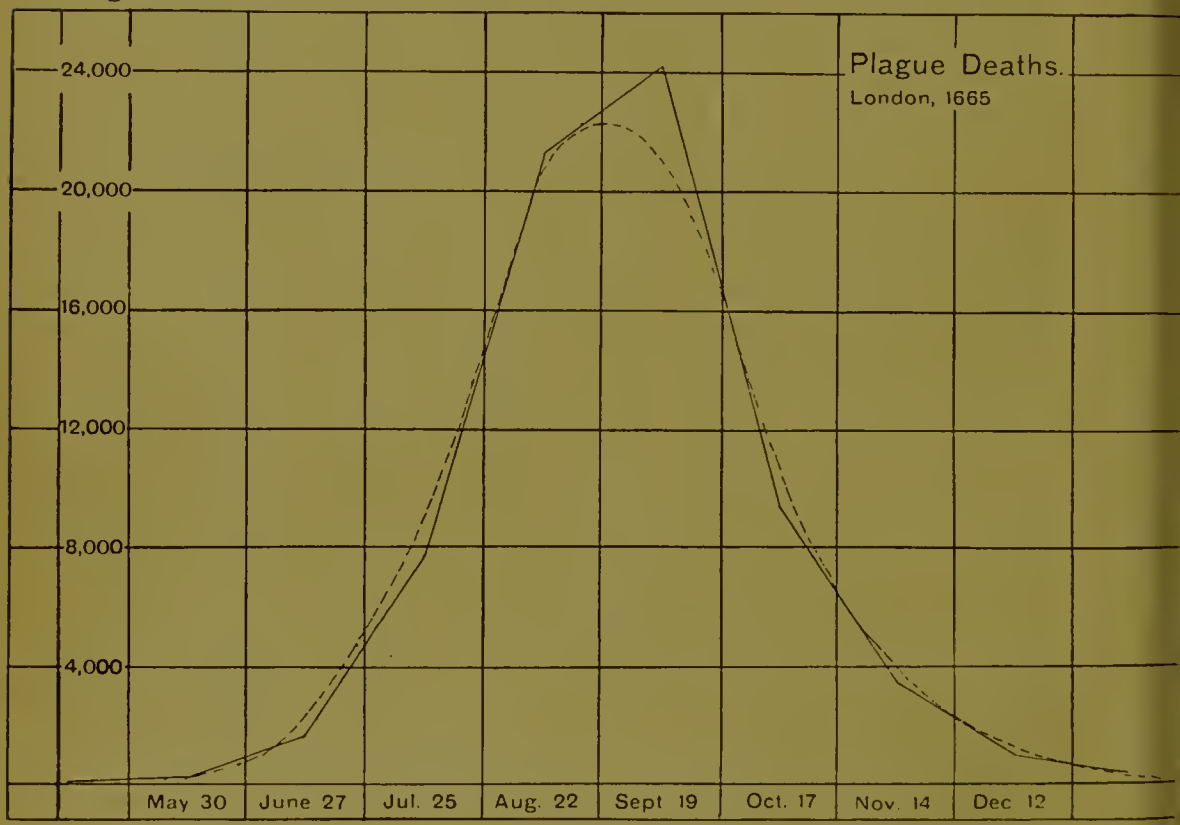
† *Loc. cit.*, vol. i. p. 305.

‡ Reports of Registrar-General for England.

§ *History of Cholera in Exeter in* 1832, Shapter, p. 208.

that in London in the spring of the same year.* The statistics of the summer epidemics in London are difficult to deal with, on account of the manner in which the death certificates have confounded the ordinary summer diarrhœa with the more deadly disease. It is also to be noted that, though the epidemics referred to occurred at different seasons of the year, the form in each disease is much the same, showing that the infecting agent is more

Diagram II.



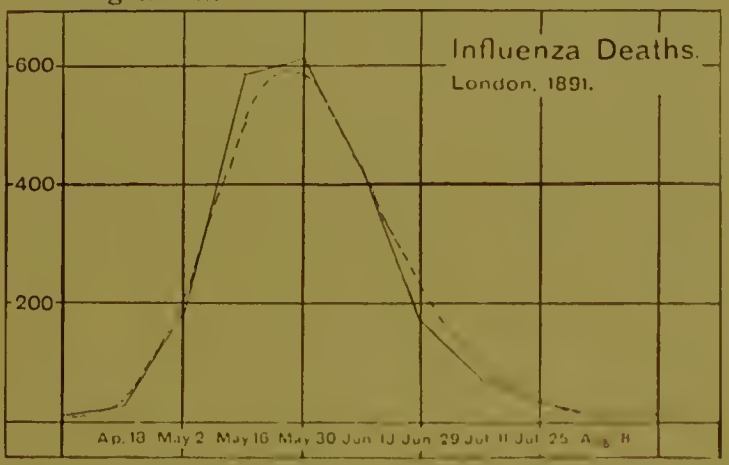
potent in determining the outburst than either season itself or seasonal constitutional differences in the population.

From the large body of figures relating to smallpox, several examples have been chosen. For the last two centuries statistics of numerous epidemics exist where the number of deaths is recorded for each succeeding week or month of the epidemic, and since 1890, when compulsory notification began, the epidemic wave can be traced for both cases and deaths.

* Report of Board of Health upon Epidemic Cholera, 1848–49, plate, p. 26.

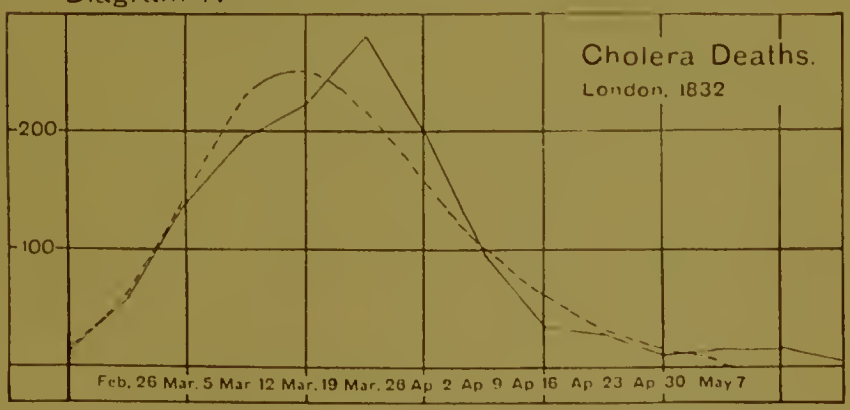
Between the conditions which existed in earlier and later times there is some difference. Then smallpox might be entirely absent from a district for long periods of time, or might recur in epidemics every three or four years. These two conditions present differences.

Diagram III



In the former, with half the city open to infection, as in Boston in 1721,* a bad epidemic might burn itself out partly from absence of material. In the latter, those susceptible would be more or less thoroughly mixed with the insusceptible, and the chance of infec-

Diagram IV

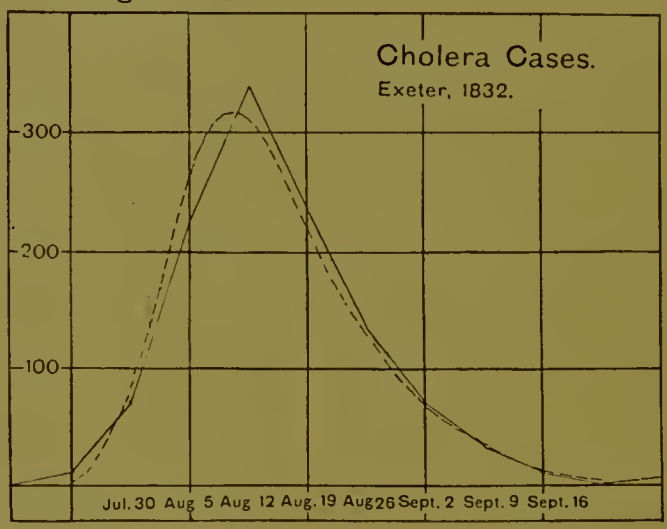


tion reduced so that the organism would have a better opportunity of producing a typical epidemic. The latter approximates more closely to the condition seen at present, where vaccination provides a large insusceptible population. Of the former, the epidemic in Boston, U.S.A., in 1721 (diagram VI., Table A, No. 9), may be

* Creighton's *History of Epidemics*, vol. ii. p. 485.

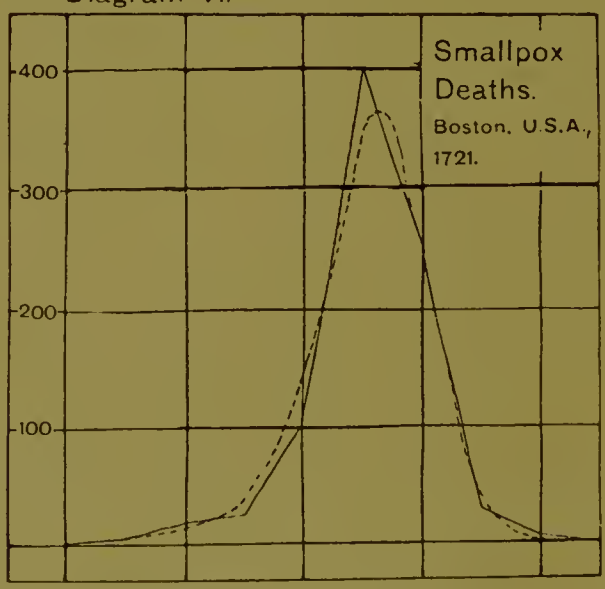
taken as an example. The population numbered at that time about 10,000, and about half only were protected by a prior attack

Diagram V.



of smallpox; of the remainder, all but 750 persons suffered from the disease. Of the latter, a large proportion were probably susceptible, though the exact conditions which were necessary to

Diagram VI.



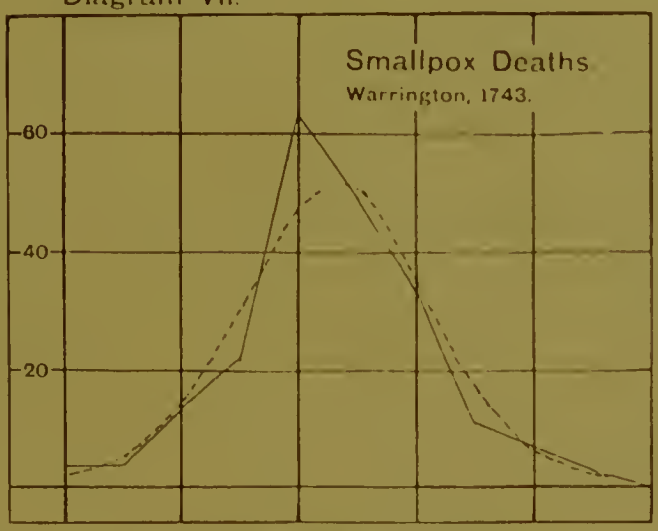
Each abscissal unit is one month.

cause their infection did not occur. In this case the decline of the epidemic was much more rapid than that usually seen,* due

* Compare Diagram XVII.

probably to the comparative exhaustion of susceptible persons. The statistics refer in this instance to the deaths alone. An epidemic occurring in Warrington * in 1743 (diagram VII., Table A, No. 8) is given for comparison. The asymmetry of this is not so great. Although both these epidemics are definitely limited as to beginning and end, the constants are still required by type IV. This great asymmetry was not the rule, however, in places where smallpox was more frequently epidemic, and where, in consequence, a sufficient dilution of the susceptible persons existed to allow the epidemic to run a more natural course, and when even at the end there were still present in the population sufficient persons open

Diagram VII.



Each abscissal unit is one month.

to infection to permit the decay in the infectivity of the organism to be observed. An epidemic in Glasgow in 1784 † is given to illustrate this (Table A, No. 10).

The epidemics of smallpox in Gloucester ‡ in 1896 and London § in 1902 (Table A, Nos. 11, 12, 13, 14) may be compared with those of last century. At first sight, in epidemics where all the machinery of modern sanitation has been brought to bear, it might be expected that the form of the course would in some way be altered. On examination, however, the course of the London epidemic is

* Report on Epidemic of Smallpox in Warrington in 1892–3, p. 7.

† Watt's *Treatise on the Chincough*, Glasg., 1813, p. 344.

‡ Report of the Royal Commission on Vaccination,—Appendix on Gloucester.

§ Reports of Metropolitan Asylum Boards, 1901–2.

seen to be very much that of presanitary days (diagrams VIII. and VIIIA.). I do not mean to infer that there is no difference in the amount of disease present in a given epidemic, but that a uniform

Diagram VIII.

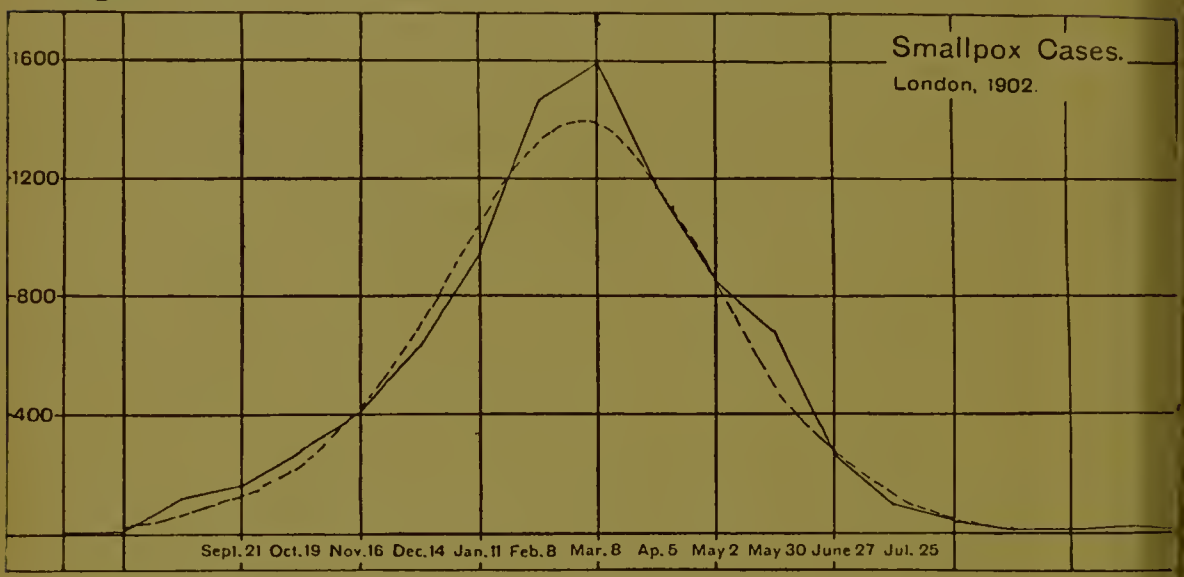
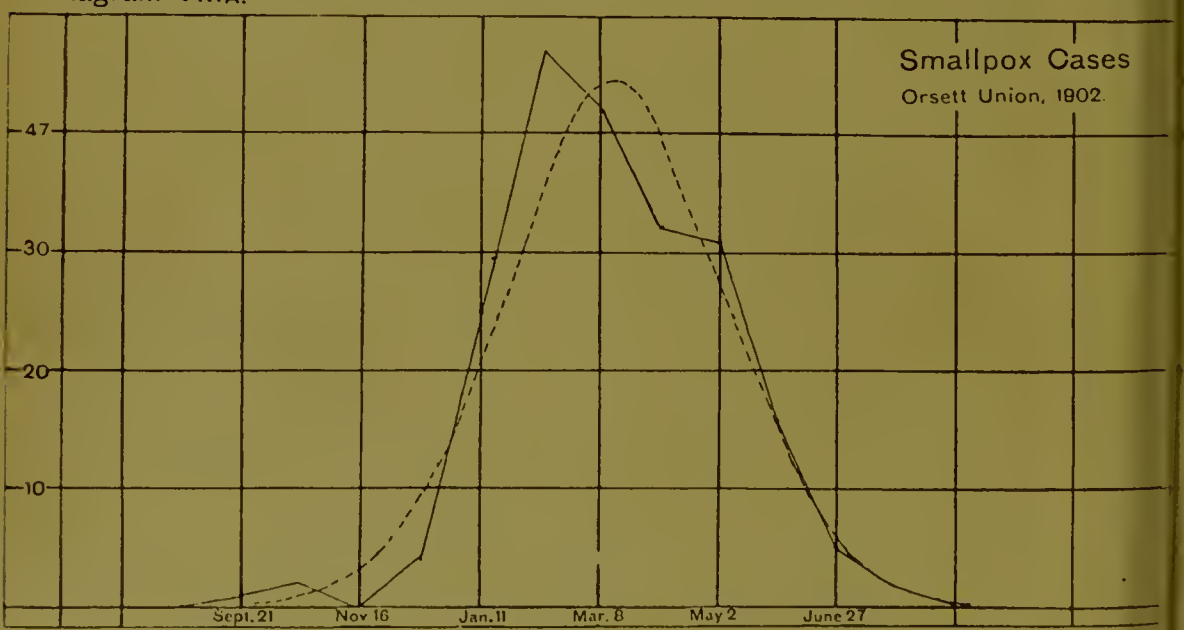


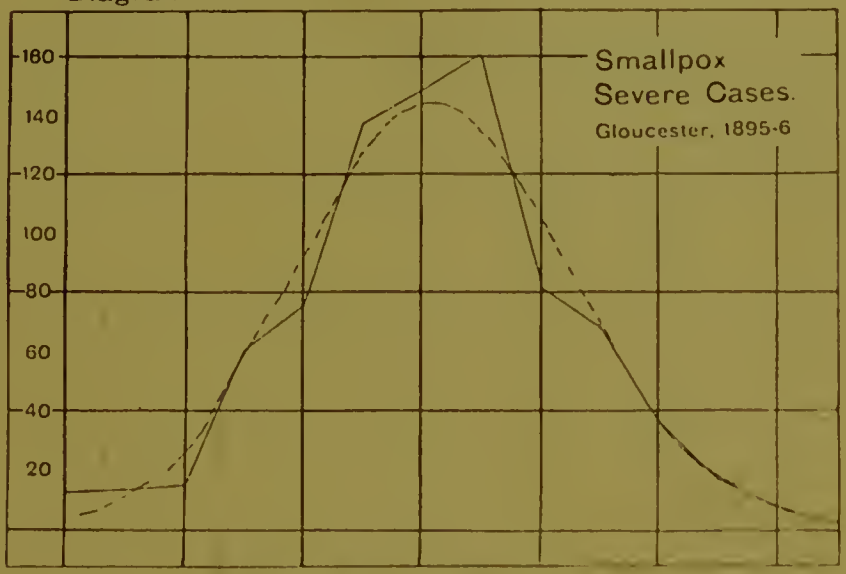
Diagram VIIIA.



force acting towards the limitation of an epidemic produces no perceptible effect on the form of the curve. In the case of Gloucester, the form of the curve has been calculated for the total cases; for the severe cases (diagram IX.); and for the

deaths (diagram X.). The general correspondence shows that the deaths may, in the case of smallpox at least, be taken as giving

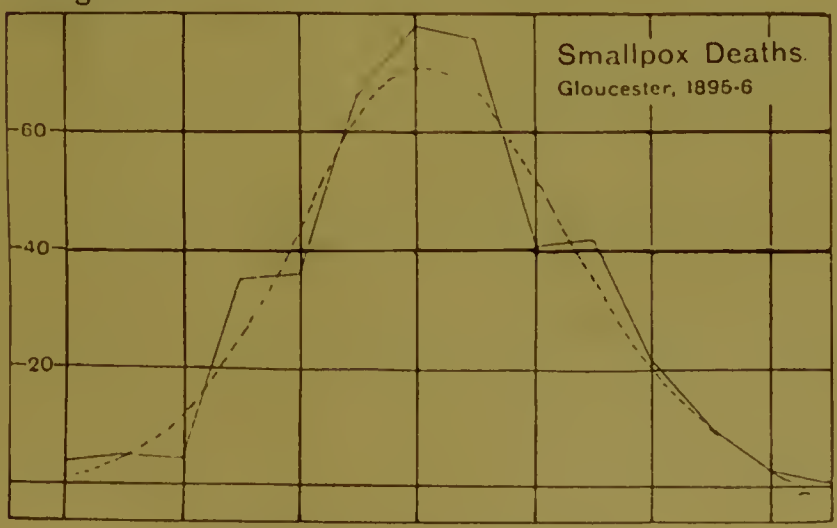
Diagram IX.



Each abscissal unit equals four weeks.

a fairly accurate representation of the course of the epidemic. With the exception of measles, this exhausts the diseases truly

Diagram X.



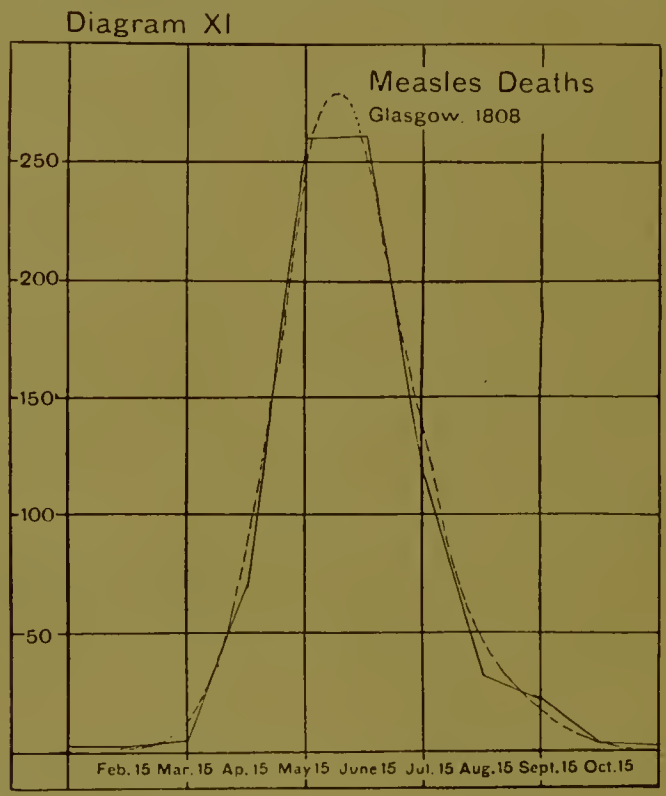
Each abscissal unit equals four weeks.

epidemic in this country. An example of the latter* in Glasgow in 1808 is given in diagram, and is seen to conform to the general

* Watt, *Treatise on the Chincough*, p. 368.

type (diagram XI., Table A, No. 18). The measles epidemic of the present day has exactly the same form.

The theory of endemic disease is more difficult. The degree of endemicity, of course, varies in different instances. Sometimes, when a disease invades a territory, there is a period of some years during which it is never absent. In some such cases we have obviously to deal with two independent epidemics, the tail of the first of which runs into the beginning of the succeeding. More



often, however, as with scarlet fever, enteric fever, malaria, etc., there is a minimum below which the amount of disease never falls. One disease, namely, zymotic diarrhœa, partakes of both characters, being truly endemic, occurring year after year at the same season, and yet in characteristic outbursts, so that it may conveniently be considered with first instance. The statistics chiefly refer to deaths, and only in a few instances to cases. The latter are in this disease specially important, as diarrhœa claims its victims mainly at the two extremes of life, and consequently nothing can be inferred *a priori* as to the relationship

of the curves representing the cases and those representing the deaths. The disease shows a marked rise at practically the same period, year after year. During the seasons at which it is nearly absent, the figures giving the weekly number of deaths are almost absolutely constant. Whether these deaths are due to the same form of diarrhœa as that which causes the summer outburst is not clear, but these numbers indicate a distribution which is essentially different from that given by the terminal portions of the probability curves. So that if the epidemic be fitted to such a distribution, some allowance must be made. The most natural assumption is to subtract a number equal to the average number of cases in the inter-epidemic period from the weekly or monthly numbers during the epidemic period.

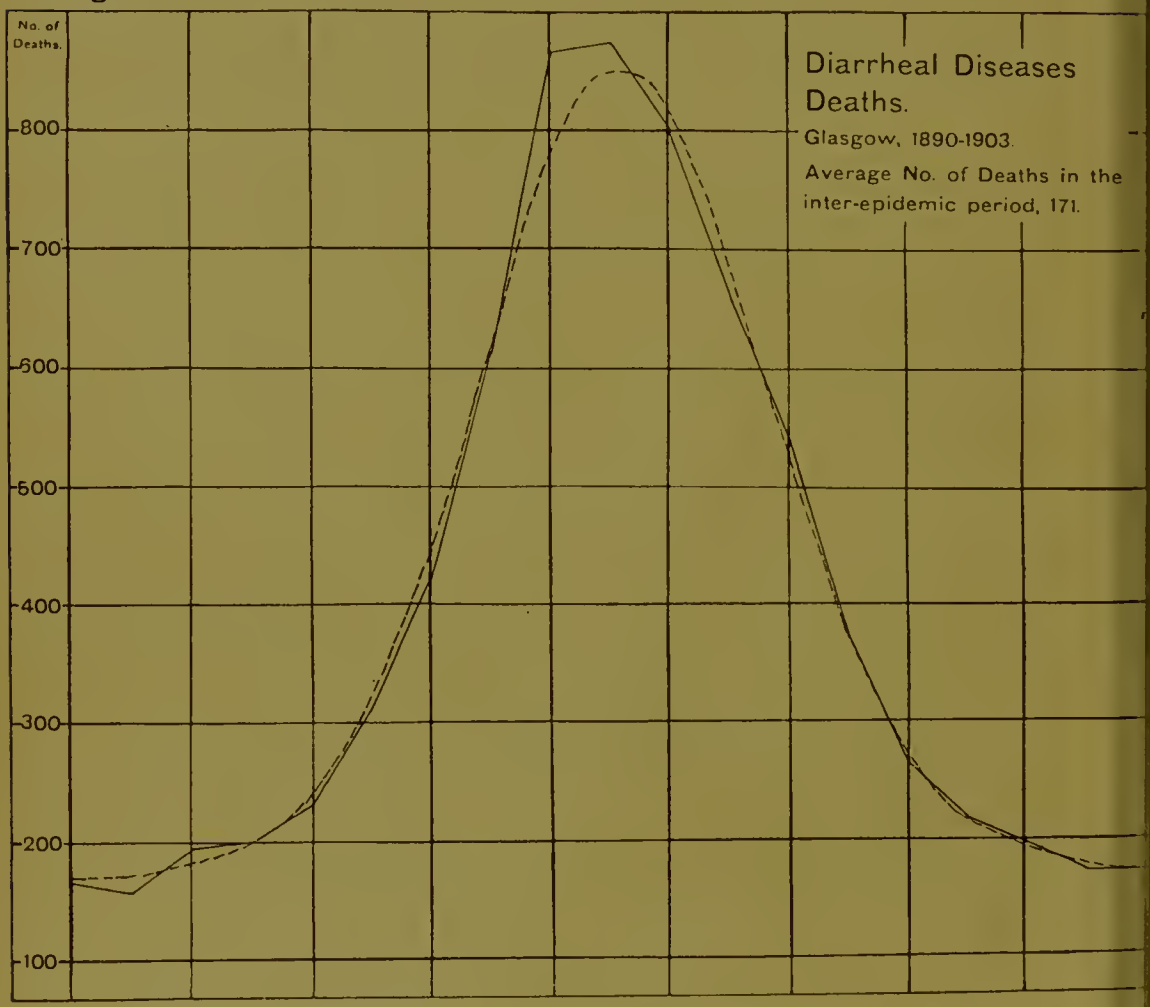
Zymotic diarrhœa forms the sole example in which an average can be obtained for the course of an epidemic, as the annual outburst lasts only about three months and has a well-defined beginning and end. As a first example the statistics of Glasgow * have been chosen, as they have been for many years less subject to these fashions of death certification which make the returns of the Registrar-General for England difficult of interpretation. The Medical Officer of Health for Glasgow has been in the habit of specially classifying these cases for himself. The result is that, while the figures may not absolutely accurately represent the true number of deaths from zymotic diarrhœa, the statistics are homogeneous. The period chosen concerns the years 1890–1903. The result is represented in the figure (diagram XII., Table A, No. 19), and is seen to be an instance of a really good fitting curve. When tested by Professor Pearson's method, the probability is found to be about ·35, which, considering the kind of case, is a good fit. As curves of cases and deaths (diagrams XIII. and XIV., Table A, Nos. 21 and 22), the figures of the Children's Hospital in Manchester † are given. Here, again, as a period of ten years is embraced in the statistics, an average is obtained. In these epidemics a curve very nearly normal in character is given by the cases of the disease, while the corresponding deaths have a dis-

* From the Notebook of the Medical Officer of Health.

† Supplement to the Report of Medical Officer of Local Government Board for 1887 on Diarrhœa and Diphtheria, p. 68.

tribution whose constants are those required by type IV. The fit of both these curves is good. At the beginning and end the divergence of the actual figures from the theoretical distribution is most marked, but it is at these points that the probable error of the statistics is very large. As an equal number was subtracted

Diagram XII.

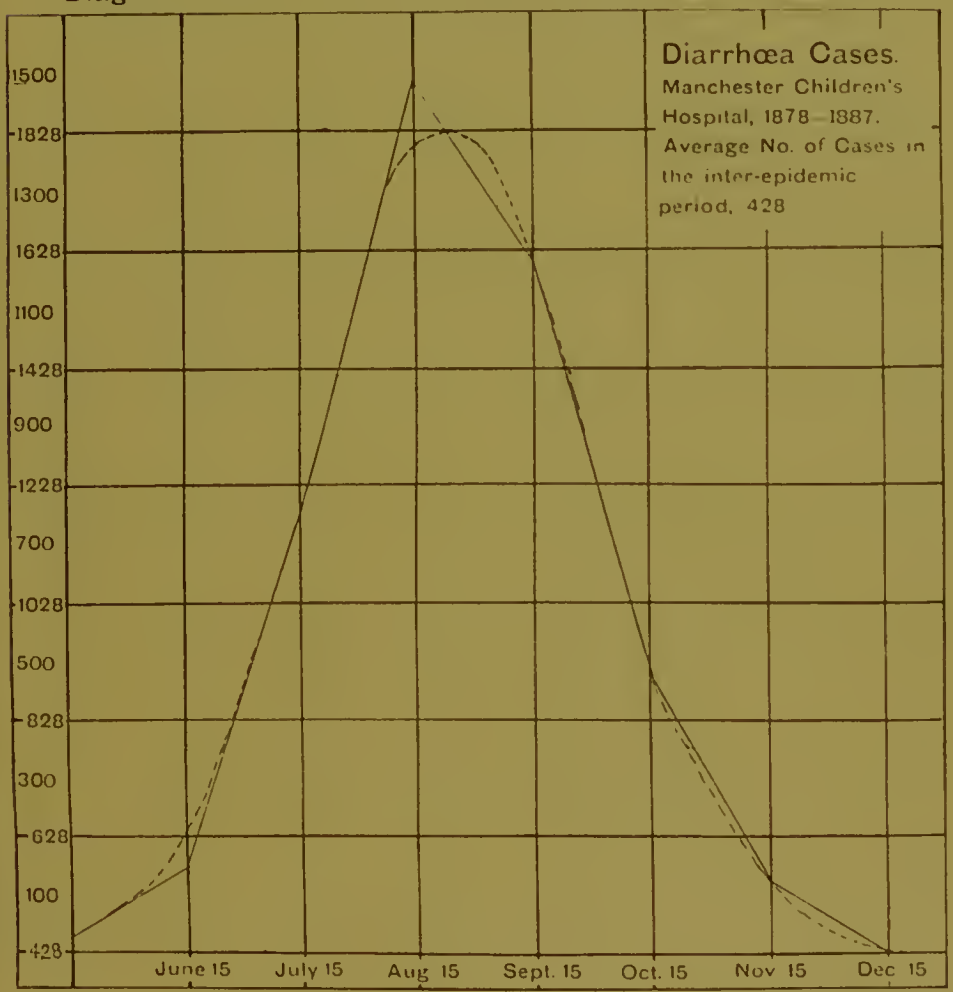


from the whole series, and as at the extremities this number, in place of being only a moderate percentage of the whole cases, becomes equal to many times the residual number of cases, it will be seen that the accuracy of the residual numbers cannot individually be very great. The curve of zymotic diarrhœa deaths, London,* for the years 1853–1903 (diagram XV., Table A, No.

* Report to Registrar-General for England, 1903.

20), is also given. The fit here is not good, but the statistics of London are not nearly so definitely homogeneous. As, however, the divergence of the actual from the theoretical is, though much greater, of the same nature as that observed in the divergence of

Diagram XIII.



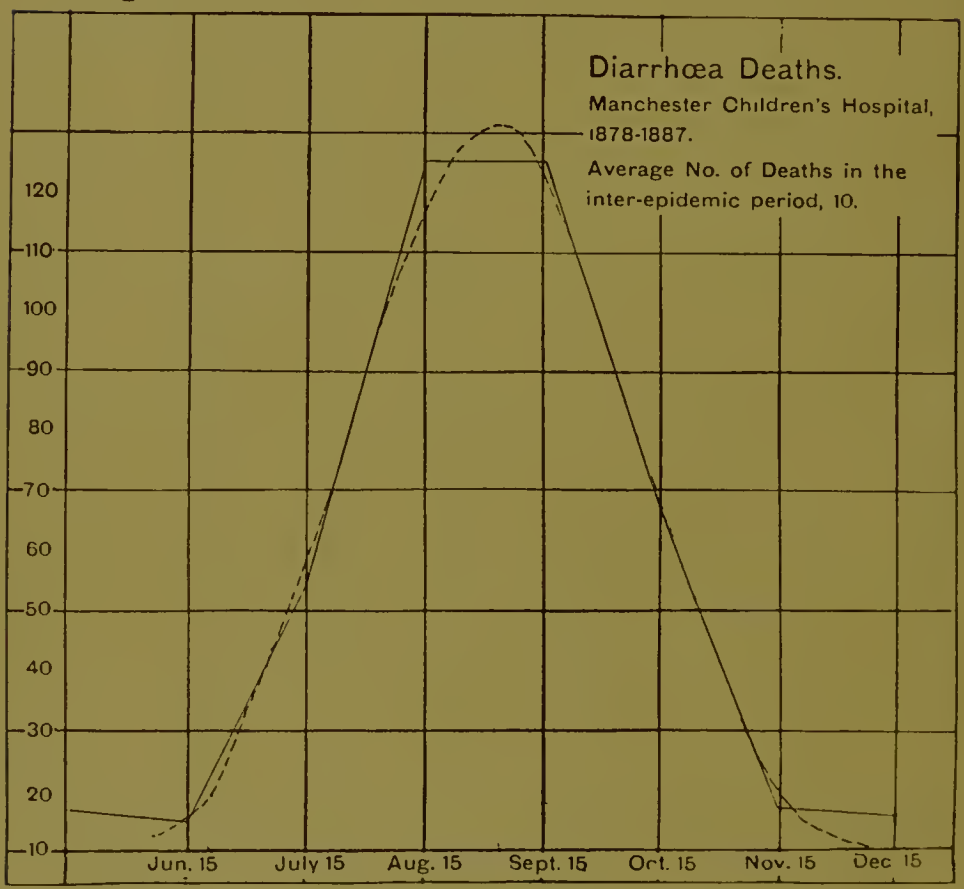
the actual and theoretical curve for Glasgow, it is possible that some other factor affects both.

So far, no exceptions have been noted to the form of epidemic distribution given by type IV. These have occurred in my work only in the case of some small epidemics of which it was difficult to give the beginning or the end, and in the case of outbreaks of endemic diseases where the tail of one epidemic runs into the beginning of another,—cases difficult to treat, as will presently be

shown. One example of this—an outburst of yellow fever * in Demerara—is given (diagram XVI., Table A, 32).

This forms the general survey of an epidemic. The applied curves seem to give a very fair approximation to the facts; but when, as in an epidemic, a disease propagates itself amid a variety of evanescent and manifestly independent influences, it is evident that, whatever the law of its spread, only an approximation can be

Diagram XIV.

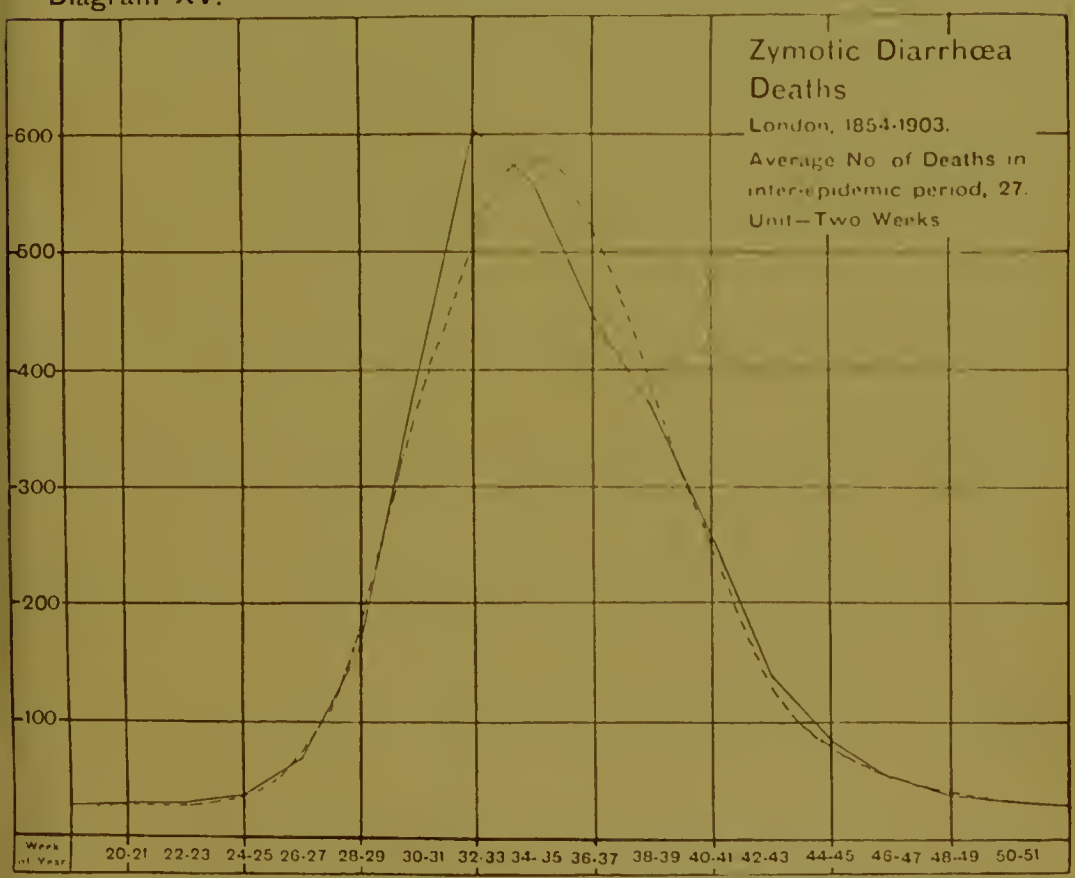


expected in the result. In the first place, there is the weather, exercising an influence which at present there is no means of measuring. Secondly, there is the distribution of the population in the infected districts. In a city, for instance, there are large areas where there is no population, or only a very sparse one, and the disease might initially appear in one of these. Thirdly, there is the nature of the housing or grouping of the people, as is seen in

* Harty, *Yellow Fever in Guiana*, London, 1820.

the case where smallpox invades a school in which there are a great number of unvaccinated children. All these factors and many other minor ones make their influence felt with more or less regularity. These, however, cannot be taken as the complete reasons for the difference noted. It is likely that there is a biological factor in addition. The infecting powers of an organism cannot be expected to obey completely regular laws of growth or

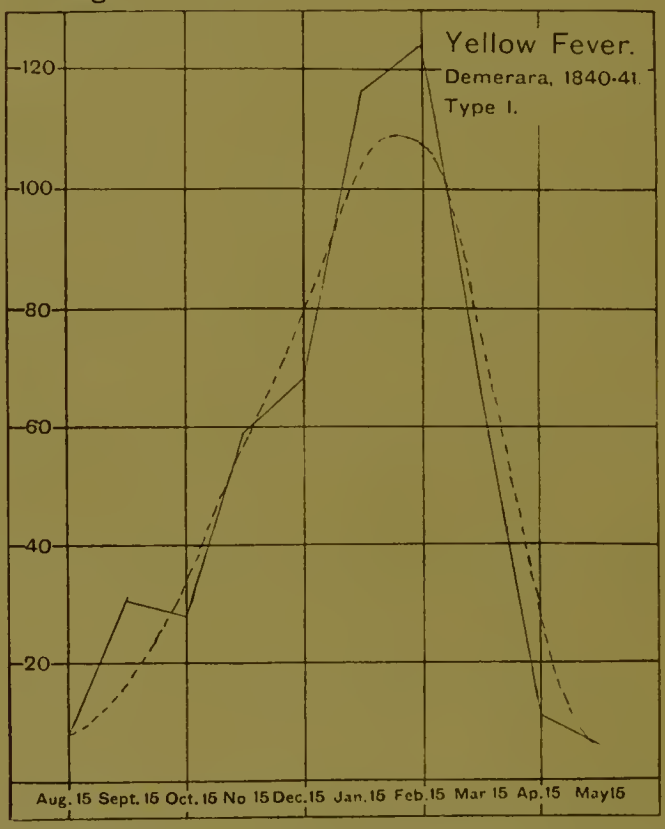
Diagram XV.



decay. As with children: some have a slow, regular growth; others make little increase in size for some years, then shoot to their full stature in a very short time. While these variations comport themselves so that in the sum the different groups adjust themselves closely to the law of error, yet in individual cases a law can only be expected to give a rough approximation, and the application of a law of averages to an individual instance cannot be expected to give good examples of curve fitting. I intend at some future time, with reference to some special disease, to make an investiga-

tion of the manner in which single epidemics vary from the average. This reason probably accounts for the larger part of the differences seen between the actual and the theoretical distributions. The striking fact is that epidemics in general hold a course whose constants with very great regularity are those of a single member of the large class of frequency distributions. It can hardly be explained on any other hypothesis than that the law which underlies the propagation of infectious diseases is such as in general to

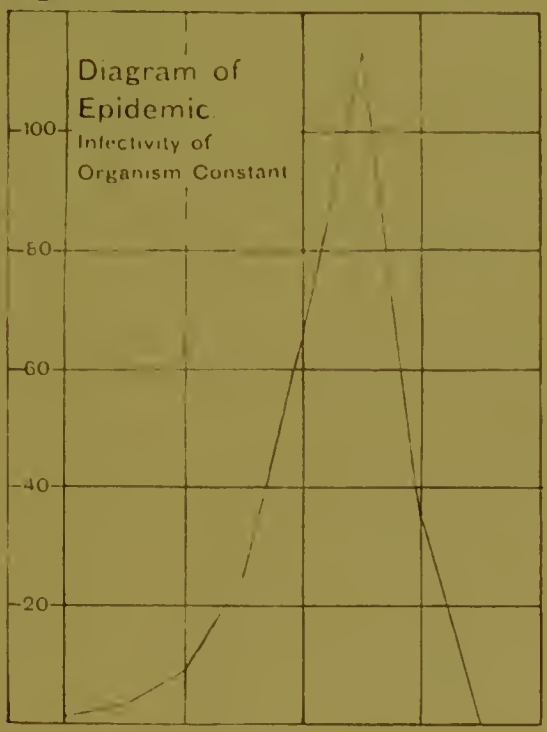
Diagram XVI.



produce such a distribution. The investigation of this is much more easily attempted *a posteriori*. The assumption that the infectivity of an organism is constant, leads to epidemic forms which have no accordance with the actual facts. If there be given a number of susceptible persons in a community, and if one, say, infect three, the whole body of the susceptible persons will become involved, and the last few remaining finally swept off. Even when allowance is made, on various hypotheses, for the chance of infection being small, because of dilution of the susceptibility

by the insusceptibility, the epidemic is only lengthened, not changed in form, and the course still represents an ascent of constantly increasing slope and a sudden drop to the original level. An example of this is given in the accompanying diagram (diagram XVII.), and it will be seen that the form bears no resemblance to any of the curves shown in the preceding pages.* Another factor is evidently necessary—and this is perhaps to be found in the loss of infecting power on the part of the organism. The rate at which this occurs must be according to some law, and it seems reasonable,

Diagram XVII.



as a hypothesis, to assume that the decline proceeds according to the terms of a geometrical progression. Thus, if at the end of

* This form is sometimes seen when a case of measles develops in a ward in which a number of susceptible children are confined. Although the sufferer be moved at once, yet the infection is present so early in the disease that usually several others succumb. In one actual case, when there were fourteen susceptible children in a ward, the epidemic developed in the following way. First *one* case; at the end of incubation period *three* cases; then a fortnight later *seven*; leaving only three to develop the disease, and these all succumbed in the next fortnight, so that the epidemic came to an end from the cause discussed and took the form of diagram XV. This, however, is under artificial conditions, and bears no resemblance to a natural outburst.

the first period of time the infecting power has declined from unity to q, say, then at the end of a second period of time it will be q^2, and so on. On this assumption, if a be the number of persons infected originally and ap the number infected by these, then the succeeding terms representing the course of the epidemic will be represented by a, $|ap$, $|ap^2q.$, $|ap^2q.\{p.q.^2\}$, $|ap^2q.\{p.q.^2\}\{p.q.^3\}|$, etc., or the general term will be given by

$$a.p^{x-1}.q^{\frac{(x-2)(x-1)}{2}}$$

which transferring to an exponential form

$$(x-1)\log p + \frac{(x-1)(x-2)}{2}\log q$$

is ae.

As q is by hypothesis less than unity, $\log q$ is necessarily negative, and in consequence the slope of the epidemic curve is seen to be that of the normal curve of frequency of error. The normal curve itself, as has been seen, occurs as an epidemic form only very rarely.

This gives an indication how the curve of an epidemic might arise, but it can hardly represent the complete solution. All that can be said is, that in general one of the curves derived by Professor Pearson to represent chance distributions makes a good interpolation formula for the ordinary course of an epidemic. These curves have been found to fit many classes of statistical grouping, and there is nothing in the method by which they are derived at all to preclude their application to this class of phenomena. These curves are the solution of the equation*

$$\frac{1}{y}\frac{dy}{dx} = \frac{-x}{a+bx+cx^2}$$

and the particular one which is found to apply to this case is that where the roots of the quadratic expression in the denominator are imaginary. Its equation is

$$y = \frac{y_0}{\left(1+\frac{x^2}{a^2}\right)} e^{-r\tan^{-1}\frac{x}{a}}$$

But though this curve expresses somewhat closely the facts of the case, yet it does not express the whole truth, as is seen when the

* See Note at end of paper.

diagrams are examined. Without exception, these show near the height of the epidemic a difference, but it is at this point alone that much divergence of the actual statistics from the interpolation formula occurs.

Although it might appear that the application of the preceding to endemic diseases was simple, yet such is not the case. Even when, from the course of an epidemic wave, it might seem easy to form equations of two succeeding epidemics, and combine them so that the sum would accurately represent the course of this epidemic, this is found to be difficult. Thus, on the introduction of a disease like plague, although the rise of the epidemic might be characteristic and the commencement of the decline also lead to the expectation that the outburst would soon approach its close, yet a recrudescence might occur before the first outbreak had finally subsided. In the case of the chief endemic diseases in this country, namely, scarlet and enteric fevers, there is a yet more difficult problem, because at no period of the year are they absent, while in the autumn epidemic outbursts occur with great regularity. Solitary epidemics are not frequent. A number, however, of the latter have been investigated, with the result that they are seen, with one exception, to conform to the same type as has been found in other diseases. For scarlet fever, for instance, the epidemic in Halifax * in 1880–1 gives the usual form (diagram XVIII., Table A, No. 27), while that of Thorshavn,† in the Faröe Islands, invaded in 1873–4 by scarlet fever for the first time for thirty years, is also seen to be of the same type (diagram XIX., Table A, No. 28). The asymmetry is, however, much greater, and the decline of the epidemic so much more gradual as to require a modification of the hypothesis that the infectivity declines according to the terms of a geometrical progression.

It would seem that with scarlet fever a considerable variability may exist in the rate at which infectivity is lost. Two epidemics of enteric fever are also given, one of which, an outbreak due to contamination of the water supply in Coventry ‡ (diagram XX.,

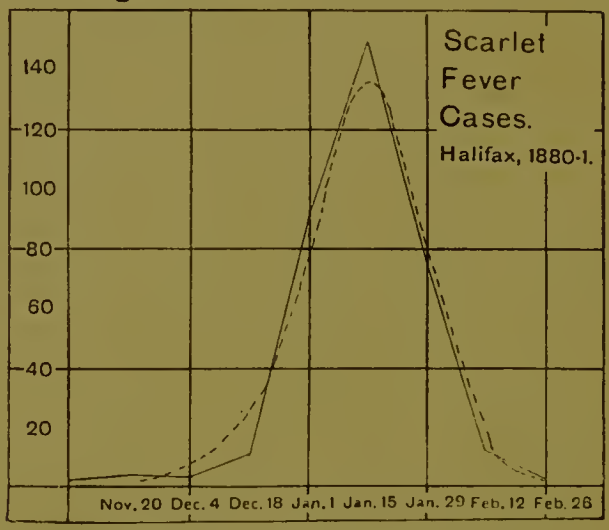
* Report of Medical Officer of Local Government Board, England, 1881, p. 60.

† Nothnagel's *Encyclopædia of Medicine*, art. "Scarlet Fever."

‡ Report of Medical Officer of Local Government Board, 1901–2.

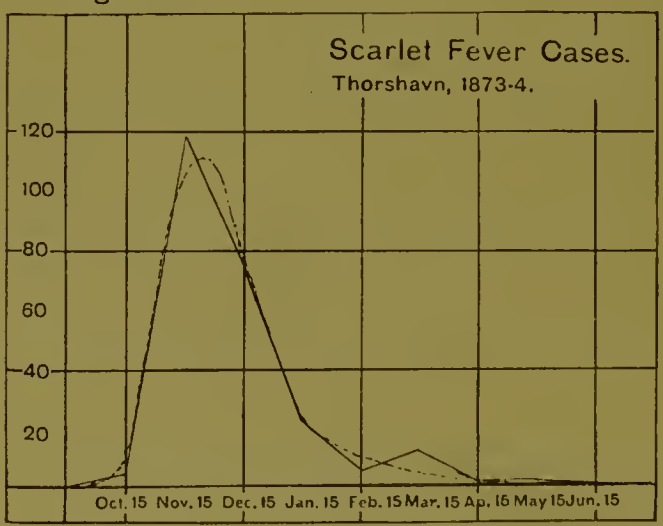
Table A, No. 25), has a similar form to that seen in scarlet fever, while the other, the great epidemic in Maidstone in 1900,* due to the same cause, but followed by large numbers of secondary cases,

Diagram XVIII.



is very asymmetrical, and takes the form of one of the chief exceptions to the conclusions arrived at (diagram XXI., Table A, No. 31). Two milk-spread epidemics of scarlet fever in Wimble-

Diagram XIX.

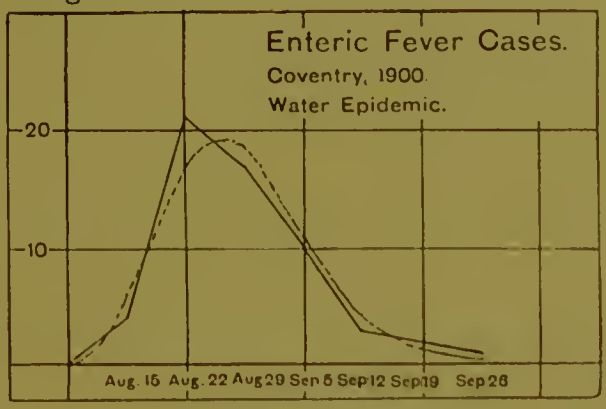


don † and Glasgow ‡ have also been investigated (diagrams XXII. and XXIII., Table A, Nos. 29 and 30). In these it will be seen

* Report of Medical Officer of Local Government Board on Maidstone Epidemic.
† *Ibid.*, 1886. ‡ Special local Report on the Epidemic, by Dr J. B. Russell.

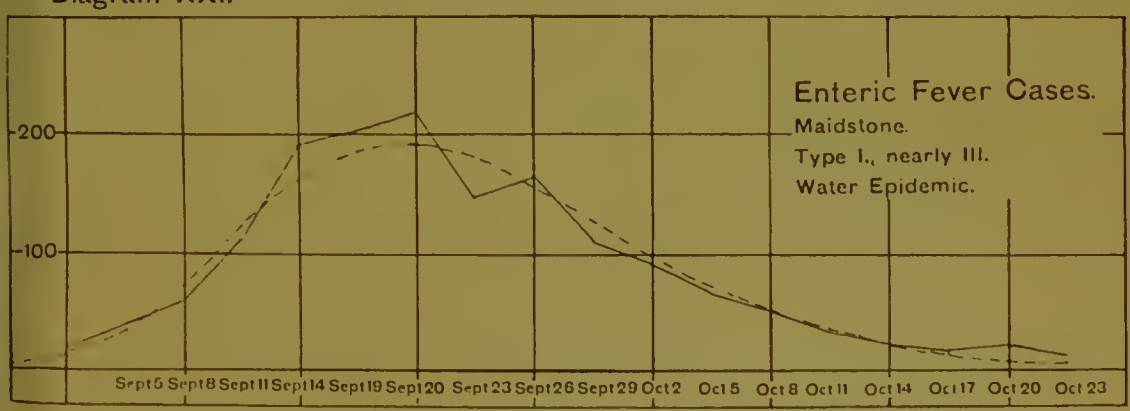
that the loss of infectivity has been specially rapid; and though a large amount of infection has been thrown into the milk, yet when it is observed that the incubation period of scarlet fever is from three to five days, it can be seen that at the time the milk supply was stopped, the organism had in both instances run

Diagram XX.



almost the complete cycle of its infectivity. The germ, therefore, of scarlet fever, though it can possibly be introduced into eruptions on the teats of cows, yet cannot evidently long maintain its infectivity if growing in that situation. When a milk supplied from such infected animals has been distributed in a new centre

Diagram XXI.

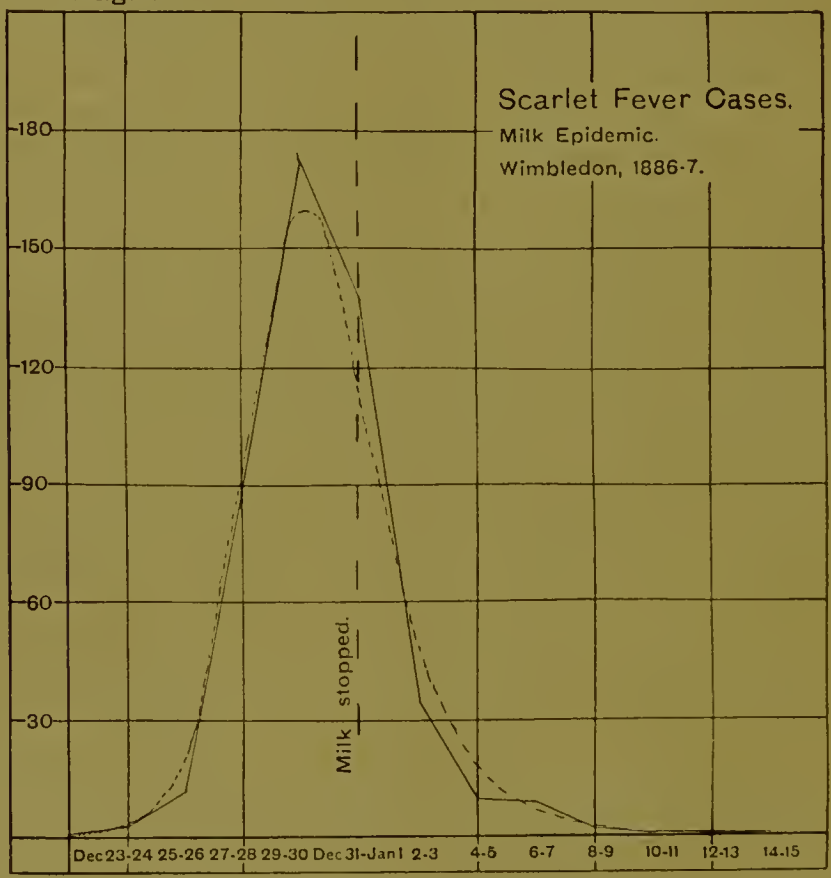


after the outburst of disease has led to its discontinuance in the infected area, though a few cases of fever might occur among the new customers, no outbreak comparable to the original one has been observed. Here a comparison may be made between the milk epidemic at Wimbledon and that previously referred to as occurring naturally at Halifax. Both curves have much the same

constants, though the former has a time unit of two days and the latter of two weeks, the same loss of infectivity taking place for each epidemic in these different periods.

The explanation of endemic prevalence which scarlet fever and enteric fever display may be sought in the fact that two factors are in action, one the true endemic prevalence, and the other the seasonal epidemic prevalence, the former varying according to a

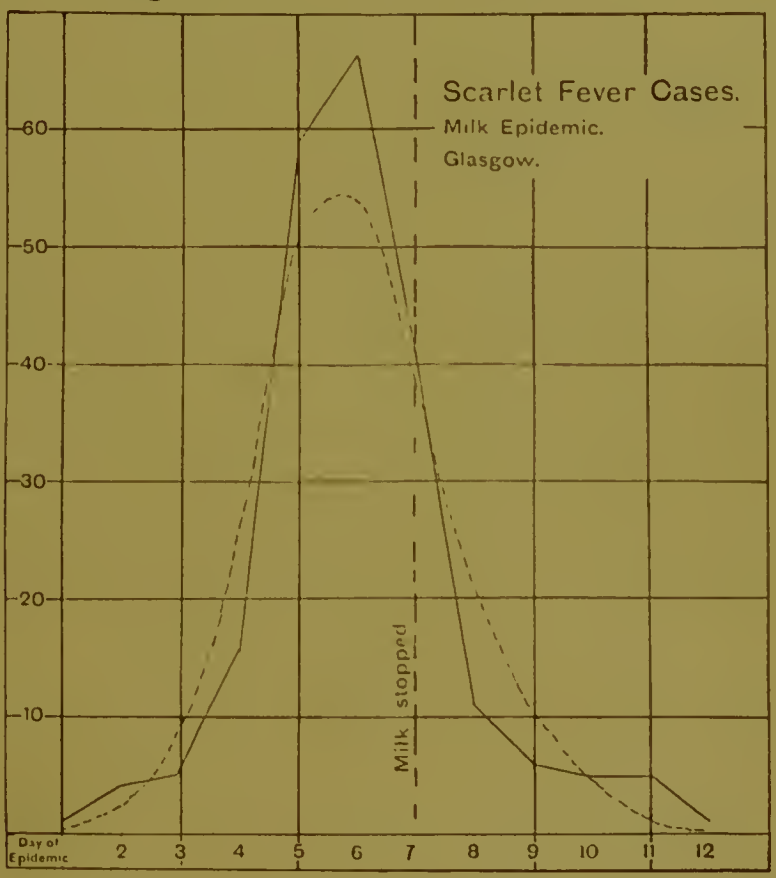
Diagram XXII.



different series of causes from the latter. There may be constantly present a certain amount of infection of a low grade which keeps up the endemic supply, while if the seasonal conditions in the early summer be suitable, some portion of this acquires a high grade of infectivity, and alone produces the true autumnal outburst. Thus a severe autumnal epidemic might readily occur though the endemic prevalence was small, and with the latter at a considerably higher level the presence of unfavourable seasonal conditions might

prevent the development of any autumnal epidemic, both of which conditions are observed to exist. I do not mean to say that the assumption of high-grade infectivity need always occur at the same season of the year—that would not be true; but in the great majority of cases it does take place in the early summer. So far

Diagram XXIII.



as my investigations go, this seems the only means of explaining the facts.

If this theory is true, certain conclusions are justified. If the number of cases or deaths be given for each week of the year, and if an average of a large number of years be made, the amount of endemic disease should be represented by a straight line, while the epidemic portion should appear on the surface of this in the form which represents the characteristic course of such an outbreak. On the other hand, on the hypothesis that the seasonal maxima and minima are due to different epidemics running into one

another, it would be expected that the average would be easily expressed by an epidemic curve, the base line of which represented the zero prevalence. In actual trial the former gives a much better representation of the facts than the latter, which seemed at first sight the more probable assumption. Two examples are given for comparison: one, the average number of deaths from scarlet fever in London * for the last thirty years; and the other, the average number of cases of enteric fever † for the last thirteen years (diagrams XXIV. and XXV.). It is seen that the fit in both cases is a comparatively good one, and is much better than any I have succeeded in forming on the basis of the second hypothesis.

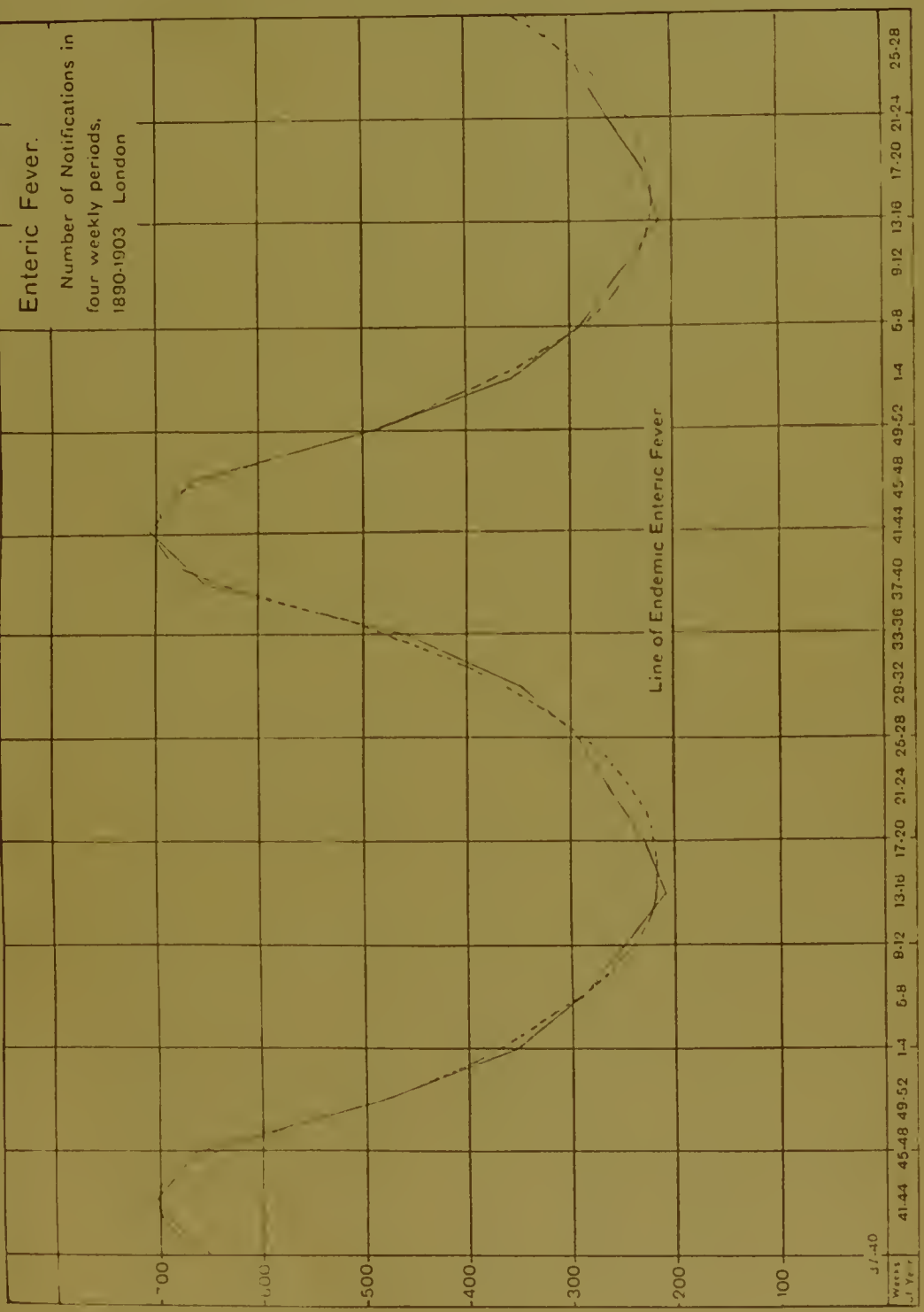
Before making any special application of the foregoing notes on the regular course of epidemics and the nature of the laws which they obey, one other point requires to be investigated. Is the distribution of an epidemic in space subject to anything like the same kind of law which regulates the distribution of the cases in regard to time? Here an answer seems easy. Given a certain amount of infection in a limited space in the midst of a uniformly distributed population, it seems natural to assume that the chance of any individual coming into the zone of infection will approximate to that given by a normal probability surface of which the maximum corresponds to the area in infection. Further, this assumption being granted, if the persons infected from this source also infect in a corresponding manner, it follows that the derived distribution will also be a normal surface, with, however, a standard deviation of a great amount.

When a disease spreads in a city, however, there are many factors which make the distribution just conjectured a form to which only an approximation can be expected. The population of a city is not equally disposed; the conditions under which people live in different districts are not identical as regards the spread of infection: especially in regard to smallpox, the amount of vaccination performed among the inhabitants of certain parts is much less than in others. Apart from these sources of error, it might be thought, as the process of spread of an epidemic is analogous to

* Reports of the Registrar-General for England.

† Reports of the Medical Officer of Health for County of London.

Diagram XXIV.



the drawing of a certain number from a limited quantity, that the

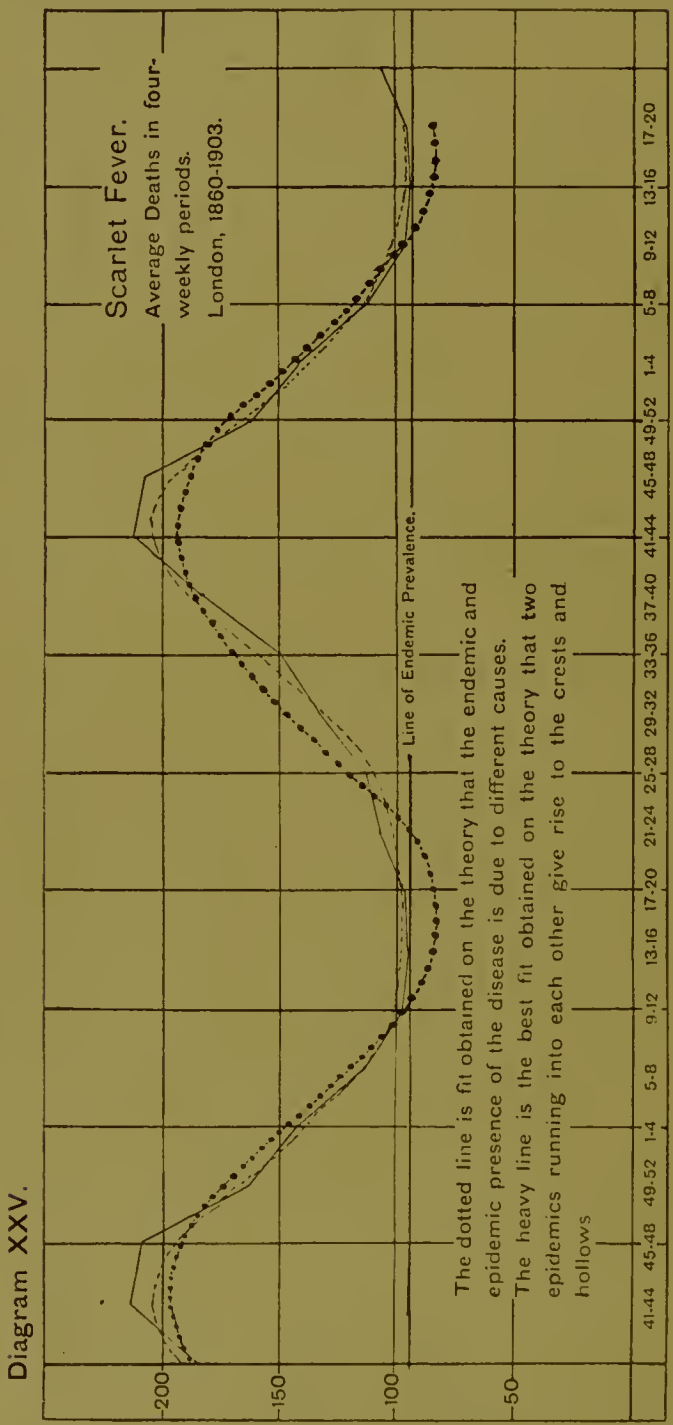


Diagram XXV.

distribution would more nearly resemble that of type IV. than that of the curve of frequency of error. This is found to be the

case in the large epidemic of smallpox in London during 1902. It is unfortunately impossible to consider the whole of this epidemic, as the centre of the outburst was in that part of London adjacent to West Ham. As the cases occurring in the latter are not included in the spot-map of the epidemic issued by the Metropolitan Asylums Board, the complete distribution is not known. If, however, London be divided into squares by a series of lines, of which one set is parallel to the boundary between London and West Ham, the distribution of the cases as these lie to the right and left of the centre of the epidemic may be studied. The summing has been taken along those parallels which are at right angles to the boundary between the two areas.*

The space distribution of the epidemic estimated in this way is seen to be of type IV., and the constants are as given in the annexed table (diagram XXVI., Table B, No. 1). The corresponding distribution of the epidemic of relapsing fever in Glasgow in December 1871 (Table B, No. 4), given for comparison, shows the same form. It does not seem necessary to elaborate evidence on this point, as the general theory is quite obvious, and the instances given sufficiently accord with it.

Two other examples will be referred to later (Table B, Nos. 3 and 4), namely, the north and south and the east and west distributions of the epidemic of smallpox in Liverpool in 1902. The latter of these is again a curve of type IV., which is nearly symmetrical, while the former is exceedingly asymmetrical, due apparently to the fact that the centre of the epidemic was adjacent to the docks, and in consequence the spread thereby so limited on one side as to prevent the development of the usual form. In this case also, however, the criterion $2\beta_2 - 3\beta_1 - 6$ is positive. The

* If the distribution were normal, it is easily seen that the sum taken in this way will also partake of the same distribution.

Thus the equation to a normal distribution is of the form

$$y = ke^{-\frac{x^2}{a^2} - b}$$

of which the integral with respect to y is

$$y = ke^{-\frac{x^2}{a^2}} \int_{-\alpha}^{\eta} e^{-\frac{y^2}{b^2}} dy$$

the curve of normal frequency.

The same holds with regard to the symmetrical form of type IV.

particular curve belongs to the class of distribution known as type VI.

We are now in a position to make a practical application of the

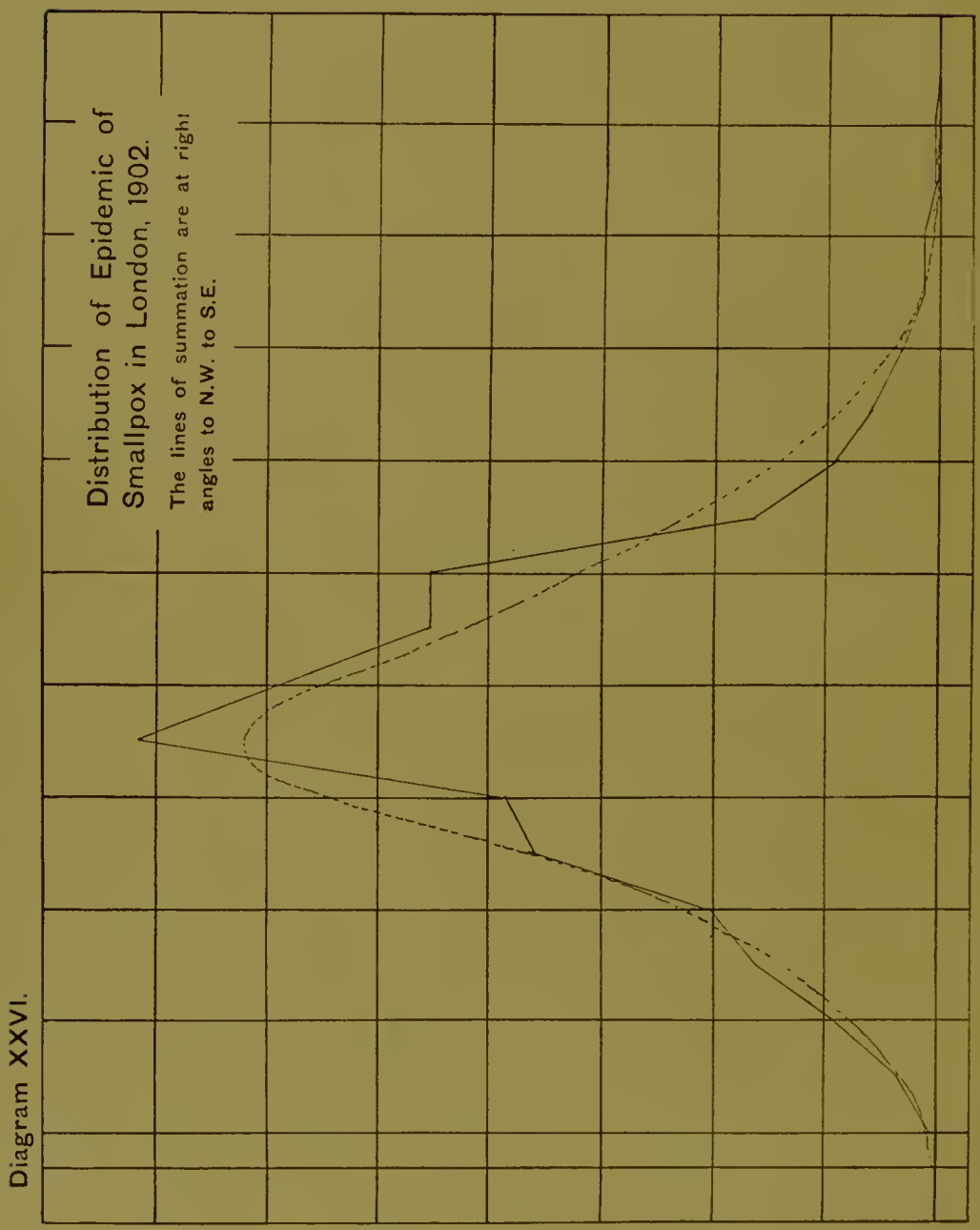


Diagram XXVI.

facts regarding epidemics which have been described in the preceding pages. The example which I have chosen is one which has a special interest at the present time, both theoretically and

practically. It concerns the method by which smallpox infection is spread from smallpox hospitals. Many believe that the infection passes aerially from the infected centre, so that in a district in the immediate vicinity of this, case after case occurs due to this cause. Now, it cannot for an instant be denied that there is often in the neighbourhood of smallpox hospitals an amount of smallpox quite out of proportion to that in the rest of the locality, but that this is necessarily due to aerial infection is a position far from tenable.

In the first place, it has been seen that a small amount of infection placed anywhere in a suitable locality is capable of producing a large epidemic; and though all the cases of the fever be promptly removed from the district, the disease may continue to spread and an epidemic develop with the typical distributions in time and space (*vide* diagrams for London, 1902). It is in relation to this type that the causes of a smallpox epidemic must be investigated.

Secondly, some centres of a town are, for reasons more or less unknown, much more suitable for the spread of an epidemic than others; and though the disease may first start in the part more or less distant from this, yet, as the epidemic proceeds and infection is introduced into this district, the cases there may so increase as to ultimately make this the chief centre of the outbreak. So that it must always be borne in mind that a local outbreak in the neighbourhood of an hospital may be as much an accident of place as a result of the proximity of the hospital.

Lastly, the form of the epidemic wave in time is of importance. If smallpox be introduced into a new district at a late period in the epidemic wave, when the organism has lost its infectivity to a certain extent, then we would, *a priori*, expect that an epidemic of shorter duration would result, unless the new locality prove fitted to temporarily rejuvenate the organism.

The epidemic in Liverpool in 1901–2 is recorded with the greatest care and detail in a report by Dr R. Reece, which was presented to the Local Government Board and published this year. This report is very suitable for the present purpose. Full details as to the dates of the occurrence of the cases, and fortnightly maps showing distribution of these cases in the city, are given. All the

fallacies from the mixing up of epidemics due to essentially different causes can thus be avoided, and the distribution of the disease in the city ascertained for any periods which may be desired. These maps also give the situation of the hospitals, and in addition are divided by parallel lines into transverse squares, and also indicate the quarter-mile zones round the different hospitals.

A careful examination of these maps shows at once that up till 31st January 1902, though there might be a suspicion that there were a few more cases in the neighbourhood of one of the hospitals, yet there was no evidence that the general distribution of the disease in the city was much affected by any but ordinary epidemic influences. The cases for this period were accordingly enumerated in the different areas given by the squares, and the form of the epidemic calculated for both the north-south and east-west distribution of the cases. It was found that the latter was of the expected type, namely, type IV., while the former, being very asymmetrical, comes under type VI. The theoretical maximum of the curves being next calculated, the centre of the epidemic—*i.e.* the locality where, up to this period, the greatest number of cases had occurred—was accurately ascertained; and with this as centre, the distribution of the epidemic in zones around it was ascertained.

It was seen at once that the epidemic, from the point of time at which it started till the 31st of March, groups itself much more naturally around the point which has been found to be the theoretical centre of the epidemic than about the hospital, and that this centre is at a distance of about one mile and a half in a direct line from the area from which infection is supposed to be disseminated. It shows at once the fallacy of calculating the prevalence of the disease in any one district and comparing with the average of the whole town, since, as we see, much the greatest number of the total cases occur in a special area, of which the zone around the hospital forms an important part.

The cases round the Priory Road Hospital, probably spread by the hospital, thus reduce to a very small proportion of the total cases which really have occurred in that zone. That this law of the distribution of the epidemic must be taken into account is easily seen when the distribution of the cases of the disease in

Liverpool is identical with that distribution seen in epidemics where there is no special local predisposing cause (London, 1902).

The incidence of cases in the neighbourhood of the Park Hill Hospital seems, however, in the Liverpool epidemic a more marked example. This case is of very special interest, as a local secondary epidemic occurred in the immediate vicinity of the hospital, beginning in a marked manner six weeks after the acute cases were admitted there. Here two points are again to be noted.

Cases of smallpox had occurred in this area prior to the opening of the hospital; and though these were few in number, there is absolutely nothing in the course of the subsequent epidemic to indicate that they were not the sufficient cause. The course in time is a typical distribution of type IV.,* which has been found to be the general epidemic form. Had special modes of infection played more than a subsidiary part in the development of this wave, there would seem to be a probability that this would not have been the case. There is, therefore, no need to assume special modes of infection. In the second place, the epidemic is so distributed in space as to have its centre three-quarters of a mile from the smallpox hospital; and if this point be joined with the centre of the hospital, it is seen that in the two quadrants which are adjacent to the hospital there are few more cases than in the quadrants which are remote. Here, again, the fallacy of neglecting the real distribution of the cases in space is capable of leading to quite untrustworthy results if the incidences in zones round the hospital are alone considered. So far as can be judged from the map, there is no greater population difference in the distribution of the population in the circle of one-quarter of a mile radius round the theoretical centre of the epidemic, and apparently little in the half-mile zone, while the latter radius includes practically the whole of the cases which strictly belong to this outburst.

Another Local Government Board report has recently been issued regarding the apparent aerial spread of smallpox in the Orsett Union District from the smallpox hospital ships in the adjacent part of the river Thames.

Here, again, when the space and time distributions are examined, there is nothing beyond the fact that the outbreak had its origin

* Table A, No. 16.

in the part of the district nearest to the hospital ships to indicate that the infection was aerial. Any kind of method which would result in the transference of the organism would be sufficient, and all gross methods are, as all who have had experience in working smallpox hospitals know, very difficult to eliminate. The course in time is illustrated in the accompanying diagram, placed for comparison below that of the corresponding epidemic in London (diagrams VIII. and VIIIA.). It will be seen that the period of maximum of the theoretical curves very closely coincides, and that the general course of the two epidemics is too much alike to require the assumption that the Orsett Union outbreak was anything but the development of an ordinary smallpox outbreak (quite possibly due originally to the smallpox ships), resulting from the introduction at a definite time of an organism of a definite infectivity.

Conclusions.

1. An epidemic is an organic phenomenon, the course of which seems to depend on the acquisition by an organism of a high grade of infectivity at the point where the epidemic starts, this infectivity being lost from that period till the end of the epidemic at a rate approaching to the terms of a geometrical progression.

2. This loss of infectivity, though realised quite clearly by Dr Farr and many other epidemiologists, has not been given the importance which is due to it. For instance, in estimating the conditions of the spread of smallpox from hospitals, it has been assumed that a constant supply of acute cases is necessary, and that, though a good number of convalescent cases are removed into a hospital, there is little risk from these, without examining whether these two factors occurred at different periods of the epidemic, when the infectivity of the organism might be greatly different. Also, in experiments such as the transmission of plague from one animal to another by means of fleas, no regard seems to have been paid to the question as to whether the organism was in a condition to transmit the infection. Attention has only been paid to whether the culture was or was not virulent—a different question altogether. Negative results attained in this manner are clearly worthless.

3. The sudden increase of infectivity in the organism points to the occurrence of some stage in its life-history at present little understood.

4. This increase may happen definitely seasonally, as in scarlet fever and enteric fever, or without apparent reason, as in the case of smallpox, which may continue smouldering in a town for a considerable period, and then suddenly give rise to an epidemic.

5. The whole explanation of the organic course of an epidemic is not to be found in this alone. Other factors, which are not clear, seem to come into play, so as to bring about differences from the form of curve to be expected mathematically on this basis.

6. That the epidemic ends because of the lack of susceptible persons has no evidence in its favour, either from the form of the curve or from the facts: *e.g.*, in the last epidemic of smallpox in London, it can hardly be believed that there were only about 8000 susceptible persons out of a population of more than 5,000,000, and that these were all confined to a small region of London.

7. And lastly, since epidemics of the same disease run pretty much the same course whether they occur in spring, summer, autumn, or winter, it would seem that the condition of the germ has much more to do with the causation of an epidemic than the constitutional peculiarity of the persons affected at the moment.

Note on the Mathematical Method employed in this Paper.

When any series of measurements are made of any natural object or phenomenon, it is in general seen that these different sizes of this object group themselves in a definite manner. These arrangements have been found by Professor Pearson capable of being represented by a series of curves distinguished as Types I., II., III., IV., V., and VI., which are the solution of the differential equation,

$$\frac{1}{y}\frac{dy}{dx} = -\frac{x}{a + bx + cx^2}$$

The solutions are as follows :—

If $b=c=0$, then

$$y=y_0e^{-\frac{x^2}{a^2}}$$

where y_0 is the ordinate at the origin. This is the normal probability curve discovered by Laplace and Gauss.

If the roots of $a+bx+cx^2$ are real,

Type I. $$y=y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2} \quad \text{where} \quad \frac{m_1}{m_2}=\frac{a_1}{a_2}$$

If both roots be real and equal,

Type II. $$y=y_0\left(1-\frac{x^2}{a^2}\right)^m$$

If one be infinite, the solution is

Type III. $$y=y_0\left(1+\frac{x}{a}\right)^{\gamma a}e^{-\gamma x}$$

If the roots are imaginary, one solution is

Type IV. $$y=y_0\frac{e^{-\nu\tan^{-1}\frac{x}{a}}}{\left(1+\frac{x^2}{a^2}\right)^m}$$

This last solution is the one which is found to be a good interpolation curve for epidemics.

Types V. and VI. do not concern this paper.

The method of fitting statistics to these curves is thoroughly described in a paper by Professor Pearson in *Biometrika*, vol. i. and vol. ii. part 1, entitled "On the Systematic Fitting of Curves to Observations and Measurements."

The theory by which the curves are derived is fully discussed in the same journal, vol. iv. parts 1 and 2, in a paper written, justifying his methods, by Professor Pearson, entitled "Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson : A Rejoinder." The subject was originally developed in two papers in the *Philosophical Transactions of the Royal Society*, and is best read in these papers. The references are "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material," *Phil. Trans.*, 1895, vol. 186A, page 343 : and "X. Supplement to a Memoir on Skew Variations," *Phil. Trans.*, 1901, vol. 197A, page 443.

The symbols used in the paper and table are as follows:— μ_2, μ_3, μ_4 represent the second, third, and fourth moments of the curve round its centre of gravity, and are obtained by multiplying each vertical strip by its corresponding abscissa, squared, cubed, or raised to the fourth power respectively, and dividing by the area of the curve.

For the determination of which curve is to be used in fitting the sign of the quantity $6 + 3\beta_1 - 2\beta_2$ is important, where $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ and $\beta_2 = \frac{\mu_4}{\mu_2^2}$. For curves of type IV. the sign of this quantity is negative.

The other constant of interest, as indicating the asymmetry of the curve, is d, which is the abscissal distance of the maximal ordinate from that through the centre of gravity.

The value of the other constants will be found in Professor Pearson's papers in the *Philosophical Transactions*

A.—Table of the Constants of the Theoretical Curves

TYPE IV.

Disease.	Locality.	Date.	Cases or Deaths.	Unit of Time.	μ_2	μ_3	μ_4	β_1
1. Miliary Fever	Oise . .	1821	Cases	1 week	7·0285	−4·0867	226·366	·0479
2. Plague	London	1665	Deaths	4 weeks	1·7758	1·3117	14·3563	·3072
3. ,,	,, .	1563	,,	..	6·1454	3·0194	123·7395	·0393
4. Cholera	,, .	1832	,,	1 week	5·2112	9·3849	120·1387	·6293
5. ,,	Exeter .	1832	Cases	2 weeks	3·5318	10·1086	98·9950	2·3411
6. Influenza	London	1891	Deaths	,,	2·4678	3·08748	27·9962	·6342
7. ,,	,, .	1891–2	,,	,,	2·3093	3·8786	30·5716	1·2217
8. Smallpox	Warrington .	1743	,,	1 month	3·1149	·7893	38·4964	·0206
9. ,,	Boston, U.S.A.	1721	,,	,,	1·3010	−1·2993	8·7323	·7656
10. ,,	Glasgow	1784	,,	,,	7·8888	2·8044	1·5082	·0161
11. ,,	Gloucester	1896	Cases	1 week	21·9069	−5·3486	1640·46	·00272
12. ,,	,, .	,,	Severe cases	4 weeks	5·7644	·1938	105·39	·00196
13. ,,	,, .	,,	Deaths	,,	5·6492	1·4905	102·115	·01322
14. ,,	London	1902	Cases	,,	6·9321	−2·4550	192·8322	·01794
15. ,,	Orsett Union	,,	,,	,,	3·3240	− ·68839	36·5115	·0129
16. ,,	Liverpool (local epidemic)	1901	,,	2 weeks	2·5540	·9005	27·8201	·0487
17. ,,	Shefficld	1887–8	,,	4 weeks	7·2097	− ·68647	174·6607	·12574
18. Measles	Glasgow	1808	Deaths	1 month	1·7039	1·13235	12·7061	·25920
19. Zymotic Diarrhœa	,,	1890–1903	,,	2 weeks	6·0870	·21329	115·4617	·00020
20. ,,	London	1854–1903	,,	,,	4·5056	2·9978	68·2995	·09825
21. ,,	Manchester	1878–1887	Cases	1 month	1·4458	·04423	6·2549	·0006
22. ,,	,, .	,,	Deaths	,,	1·62106	− ·16702	10·8032	·006
23. ,,	Islington	1857–62	Cases ages (1–5)	4 weeks	2·3797	·47904	21·7134	·01703
24. ,,	,, .	,,	Cases ages (5–)	,,	2·10781	− ·90387	19·7548	·08724
25. Enteric Fever	Coventry	1900	Cases	1 week	1·76744	1·95543	11·71872	·3471
26. ,,	Rotherham	1892	,,	,,	5·27047	5·14338	94·1654	·17811
27. Scarlet Fever	Halifax	1880–1	,,	2 weeks	1·35199	−1·13994	3·55176	·52582
28. ,,	Thorshavn	1873–4	,,	1 month	1·83684	3·28100	22·22123	1·7371
29. ,,	Glasgow	.	,,	1 day	3·08965	3·18762	43·03616	·34451
30. ,,	Wimbledon	.	,,	2 days	1·92734	2·60173	26·30963	·94548

TYPE I.

					μ_2	μ_3	μ_4	β_1
31. Enteric Fever	Maidstone	1900	Cases	3 days	12·0281	26·7769	472·168	·4121
32. Yellow Fever	Demerara	1840–1	,,	1 month	3·6612	−3·0409	37·5459	·1883
33. Zymotic Diarrhœa	Islington . .	1857–62	Cases under one year	,,	1·45193	·57488	4·61391	·10797

B.—Table giving the Constants of the Theoretical Curves

1. Smallpox	London	N.W. to S.E.	Cases	..	7·4279	7·6882	196·6696	·1442
2. ,,	Liverpool	E. & W.	,,	..	4·1247	2·6121	66·1562	·0972
3. ,,	,,	N. & S.	,,	..	1·6609	2·1135	13·0490	·9771
4. Relapsing Fever	Glasgow	E. & W.	,,	..	1·5224	·2124	9·7323	·01025

CORRESPONDING TO THE COURSES OF EPIDEMICS.

TYPE IV.

β_2	d	md	r	ν	a		No of Diagram.	Remarks.
4·583	·14798	·6527	6·8217	·9470	5·0493	..	I.	
4·5526	·2346	1·2859	8·9635	3·2622	3·5332	..	II.	
4·2209	..	..	8·2141	..	..	..	..	
4·4240	·71213	6·8129	17·1334	18·931	6·1659	..	IV.	
7·9384	·9448	5·5113	9·6667	105·175	·50655	..	V.	
4·5863	·4685	3·7369	13·9478	11·7934	4·4195	..	III.	These epidemics are almost identical although the max. of the first was in May and the second in January.
5·7329	·5946	4·0726	11·6988	..	..	..	..	
3·9677	·0824	·4712	9·4378	·8711	5·1049	..	VII.	
5·1578	– ·3342	1·3523	10·0932	4·0054	2·7823	..	VI.	
3·6593	·1530	1·2434	14·2556	1·7394	10·1675	.	..	
3·4183	– ·0905	·8825	17·4959	·8133	18·986	..	..	
3·1718	·0479	·9715	38·551	2·5513	14·601	..	IX.	
3·1997	·1181	2·2574	36·2085	5·8743	13·94	..	X.	
4·01274	– ·11288	·62727	9·1136	·7625	7·4959	..	VIII.	
3·3040	– ·0897	1·17172	24·1147	3·2589	8·6703	..	VIIIA.	
4·2688	..	..	..	..	..	..		
3·35995	·4328	..	..	..	..	..	..	
4·3664	– ·2172	1·25287	9·53494	3·55635	3·35904	..	XI.	
3·11622	·01688	·47918	54·7740	1·4512	18·0862	.	XII.	
3·86141	·29294	4·87785	31·3251	..	10·6226	..	XV.	
2·99058	..	..	..	..	..	..	XIII.	Normal curve.
4·1111	– ·0318	·1663	8·4564	·405896	3·4727	..	XIV.	
3·8333	·06834	·42576	10·45943	·76214	4·7257	..	..	
4·44641	– ·12539	·60384	7·6313	1·2489	3·6896	..	..	
3·7514	·34452	5·7297	31·26136	3·94458	4·5408	..	XX.	
3·38995	..	..	..	..	..	..	..	
5·2256	·2482	1·2067	7·7248	3·3742	2·7633	..	XVIII.	
6·5863	·5151	3·5450	11·7664	4·8054	·86803	..	XIX.	
4·5083	·34823	2·01485	9·5718	4·05043	4·76246	..	XXIII.	
7·08272	·32812	1·27706	5·78413	2·68123	2·75496	..	XXII.	

TYPE I.

β_2	r	m_1	m_2	a_1	a_2	d	
3·2638	15·6768	..	..	..	..	..	XXI.
2·7998	10·0166	2·1776	5·8390	3·708	9·942	·6224	XVI.
2·7010	3·33101	1·7606	6·6703	1·7773	6·7345	·5824	..

CORRESPONDING TO THE DISTRIBUTION OF EPIDEMICS IN CITIES.

	r	d	md	ν	a			
3·5045	20·8514	·4269	4·8780	9·1465	11·1294	.	XXVI.	Type IV.
3·8885	7·3258	..		..	..	.	..	Type IV.
4·7303	..	..		..	..	..	..	Type VI.
4·0435	8·8507	·0440	·2388	·6127	3·451	..	..	Type IV.